Choosing a Research Question: Implications for Efficient Clinical Trials

John H. Powers, MD
Senior Medical Scientist
Leidos Biomedical Research in support of
Clinical Collaborative Research Branch
NIAID/NIH

Introduction

- Choosing a research question as first step in research endeavors
  - Types of research questions
  - Process of working from one question to the next – building a research "portfolio"
- Choice of research question and implications for design of studies
  - Influence on study design and sample size
  - Most efficient use of resources

2

Boundaries between Research and Practice

- Clinical research – designates activity designed to test hypothesis, permit conclusions, and thereby develop or contribute to generalizable knowledge (expressed in theories, principles, and statements of relationships)
- Clinical practice – interventions designed solely to enhance well-being of individual patient or client and have reasonable expectation of success, purpose to provide diagnosis, preventive treatment or therapy to particular individuals
  - Belmont Report on Ethics in Human Experimentation

3

Scientific and Ethics

- Study that cannot contribute to generalizable knowledge is not ethical
- Puts patients at risk of harm (even from minor inconvenience) for no benefit to anyone
- Scientific validity is not a "nice to have" but a requirement of all research
- Validity = ability of study to correctly answer research question posed

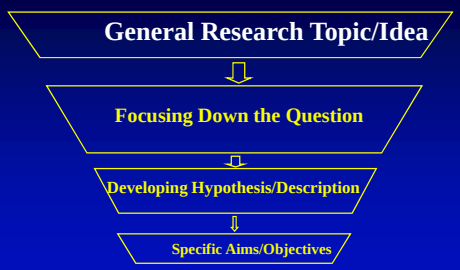
4

Science

- *"It is not what the man of science believes that distinguishes him, but how and why he believes it. His beliefs are tentative, not dogmatic; they are based on evidence, not on authority or intuition."*
- Bertrand Russell, *A History of Western Philosophy*, p. 527.

5

Focusing the Question



6

Choosing a Topic

- **Idea come from life experiences**
  - **Clinical practice**
    - **"What do I have" (diagnosis)**
    - **"How bad is it" (prognosis)**
    - **"Can I give it to my family?" (natural history)**
    - **"How did I get this?" (etiology)**
    - **"Will this stuff help me?" (treatment or prevention)**
- **Clinical research often focuses on biology but can focus on other areas**
  - **"Why do people come to the doctor for disease X?"**
  - **"What behaviors influence outcome in disease Y?"**

7

Choosing a Topic

- **Developing better tools to evaluate and describe disease is good place to start**
- **Need to characterize natural history of disease in order to study it**
- **Need to have valid measures of outcomes in order to 1) measure them and 2) describe risk factors for outcomes**
- **Example of measurement tools:**
  - **Seriousness/severity of disease (comparing baseline factors with outcomes)**
  - **Developing outcome measures (measures of morbidity e.g. functional status and symptoms)**

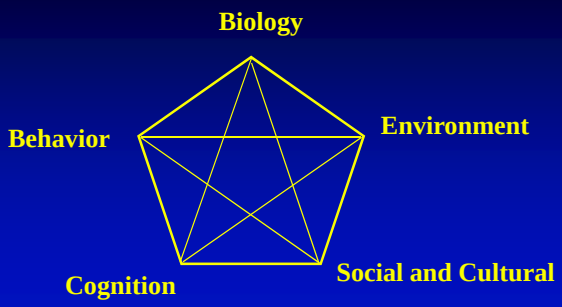
8

Descriptive Research

- *"Classification is fundamental to the quantitative study of any phenomenon. It is recognized as the basis of all scientific generalization and is therefore an essential element in statistical methodology. Uniform definitions and uniform systems of classification are prerequisites in the advancement of scientific knowledge. In the study of illness and death, therefore, a standard classification of disease and injury for statistical purposes is essential."*
- *Manual of the International Statistical Classification of Diseases, Injuries, and Causes of Death* (ICD) 1957 (Introduction, pp. vii-ix)

9

Areas for Investigation
Many Related Areas Other than Biology



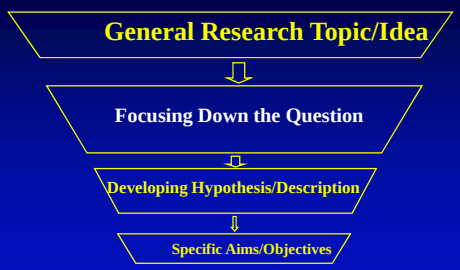
10

Choose a Broad Topic

- Choose an idea that is interesting – you have a lot of work to do!
- Choose a topic that is timely and relevant (need to answer the "so what" question) – a question worth answering
- Choose a topic that is answerable – keep in mind time and resource constraints
- F.I.N.E.R. (Hully and Cummings, Designing Clinical Research 1998)
  - Feasible – BUT feasible and invalid = unethical
  - Interesting
  - Novel
  - Ethical
  - Relevant
- Mentor can help with all these facets of project

11

Focusing the Question



12

Focusing the Question

- Need to evaluate the medical literature and other sources to evaluate current knowledge in research area
- In many areas, much of what we know is less certain than generally believed
- Review and guidelines are places to start, but represent synthesis of others views on the data – need to review for yourself

13

Why Most Published Research Findings Are False

John P. A. Ioannidis

Summary

There is increasing concern that most current published research findings are false. The probability that a research claim is true may depend on study power and bias, the number of other studies on the same question, and, importantly, the ratio of true to no relationships among the relationships probed in each scientific field. In this framework, a research finding is less likely to be true when the studies conducted in a field are smaller; when effect sizes are smaller; when there is a greater number and lesser preselection of tested relationships; where there is greater flexibility in designs, definitions, outcomes, and analytical modes; when there is greater financial and other interest and prejudice; and when more teams are involved in a scientific field in chase of statistical significance. Simulations show that for most study designs and settings, it is more likely for a research claim to be false than true. Moreover, for many current scientific fields, claimed research findings may often be simply accurate measures of the prevailing bias. In this essay, I discuss the implications of these problems for the conduct and interpretation of research.

factors that influence this problem and some corollaries thereof.

Modeling the Framework for False Positive Findings

Several methodologists have

is characteristic of the field and ca vary a lot depending on whether t field targets highly likely relationsh or searches for only one or a few true relationships among thousan and millions of hypotheses that m

finding has been claimed based o achieving formal statistical signific the post-study probability that it is is the positive predictive value, PP The PPV is also the complementa probability of what Wacholder et a have called the false positive repo probability [10]. According to the

should be interpreted based only on *p*-values. Research findings are defined here as any relationship reaching formal statistical significance, e.g., effective interventions, informative predictors, risk factors, or associations. "Negative" research is also very useful.

Moreover, for many current scientific fields, claimed research findings may often be simply accurate measures of the prevailing bias. In this essay, I discuss the

Ioannidis JP PLoS Medicine 2005;2(8):e124

What Do We Know Already?
The "Knowledge Gap"

1. How do you define the problem – clinical features, signs, symptoms, laboratory values
2. How do you diagnose the problem? What is impact of diagnosis on outcome?
3. How large is the problem? Magnitude of problem in terms of number and types of persons affected
4. What is impact of problem? Attributable morbidity and mortality (often assumed)
5. What are risk factors for getting problem? (not necessarily causal)

15

What Do We Know Already?
The "Knowledge Gap"

6. What is prognosis of problem? Morbidity and mortality

7. What factors modify prognosis independent of treatment? (Confounders)

8. What interventions can mitigate problem? Interventions can include drugs, devices, biologics, and behaviors

9. What factors are effect modifiers of treatment? Effect size differs depending on presence of factor

10. How do interventions compare to each other?

16

What Do We Really Know?

"The greatest obstacle to discovery is not ignorance, it is the illusion of knowledge"

- Daniel Boorstin

17

What Do We Know Already?
The "Knowledge Gap"

- What do we know about previous questions?
  - Often challenging to convince others we don't really know what we think we know
  - *Argumentum ad verecundiam* fallacy (argument from authority)
- What is quality of data? Validity, reliability and precision, biases in previous data
- Has evidence been independently confirmed?
- Is evidence consistent across different populations? Is a difference biologically plausible?
- For interventions: dose, duration of therapy, combination therapy

18

Overall Research Plan

- After evaluating what is known, need to focus on one part of research gap, one piece of the puzzle
- How does this research fit into an overall plan? What happens after this study (by you or someone else)?
- A single study cannot possibly answer all questions about a topic
- Research fits into an overall model/theory of a problem

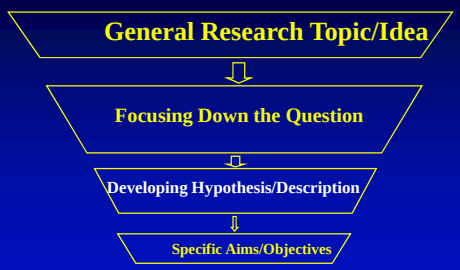
19

Feasibility

- What kinds of information do you need to answer the question?
  - Population – both test and control groups in analytical study
  - Exposures
  - Outcomes
- What kinds of information are available?
- What resources are needed to obtain data needed?
- Is there access to resources needed?
- Feasibility does NOT mean using invalid methods because that is "the best that can be done"

20

Focusing the Question



21

Developing Hypothesis or Description

- **Hypothesis – a statement about what investigator believes to be true about nature and relationships of two or more variable to each other**
- **Hypothesis testing entails a comparison, but not all research is comparative**
- **Differentiate *qualitative* and *quantitative* research**
- **Differentiate *descriptive* from *analytical* research**

22

Developing Hypotheses
Qualitative and Quantitative Research

- **Qualitative**
  - **Aim is complete detailed description – words**
  - **Develop observations for further testing**
  - **Only know roughly in advance what to look for**
  - **Early phase of research project**
  - **Design emerges as study unfolds**
  - **Researcher is data instrument**
- **Quantitative**
  - **Aim to classify features, count them - numbers**
  - **Construct statistical models to explain observations**
  - **Clearly state in advance what to look for**
  - **Later phases of research project**
  - **All aspects carefully designed before data collection**
  - **Researcher uses tools (questionnaires, equipment) to collect numerical data**

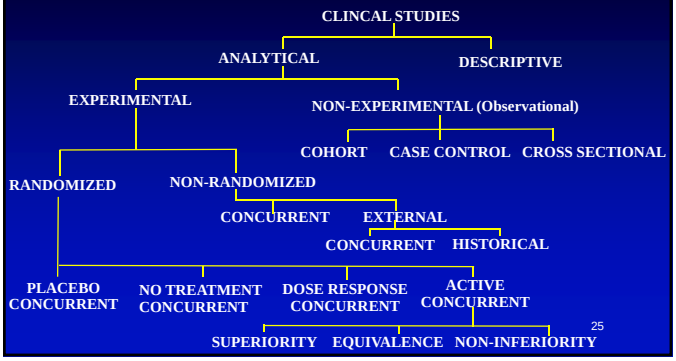
23

Developing Hypotheses
Descriptive and Analytical Research

- **Descriptive research – provide an account and delineate components of a problem**
  - **Case report**
  - **Case series – more data does not necessarily increase validity**
- **Analytical research – testing one of more hypotheses in a quantitative fashion**
- **Distinction not as clear as descriptive research often contains comparisons (but cannot assess causality) and analytical research often contains descriptions**
- **Push for "hypothesis driven" research tends to make descriptions sound less valuable but descriptions help form hypotheses**

24

Choosing A Design
Types of Clinical Studies



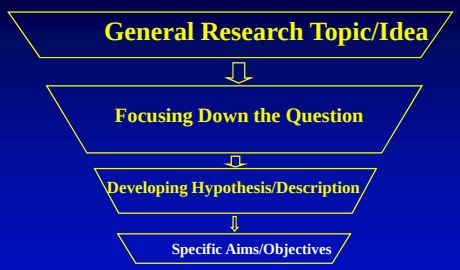
25

Examples of Hypotheses

- The more specific the better:
- " Antibiotics are effective in acute otitis media in children"
- "Amoxicillin is effective in acute otitis media in children who are between 2 and 6 years of age"
- "Amoxicillin is effective compared to placebo in reducing pain in children ages 2 to 6 years with initial episodes of acute otitis media"

26

Focusing the Question



27

Specific Aims and Objectives

- Choosing an overall research questions gives you a "why" (the rationale for doing the study)
- Next need to answer the questions related to specific measurements and define them:
  1. Who – define population under study
  2. Where – setting in which study will occur
  3. When –
     - what time frame of analysis; "from January 2000 to January 2009"
     - Prospective or retrospective – hypothesis in relation to data, not how collected
  4. What – variables of exposure, intervention and outcome (content validity)
  5. How – what tools to use to measure variables (construct and criterion validity)
     - Need to be as specific as possible regarding measurement variables and how to measure them
     - Failing to plan is planning to fail
     - Avoid circular or vague language: "Clinical outcomes will be divides into clinical success and failure"

28

Right Tools for the Job

29

Common Pitfalls

- Letting "feasibility" issues change the question to one not worth answering or answering it in an invalid way
- Taking on too many questions and thereby answering none
- Lack of clarity of hypothesis and lack of clarity on study design
- Vague specific aims and variables and unclear measurement properties of tools

30

Designing Trials Efficiently

31

Definitions

- **Clinical trial – a controlled prospective study enrolling human subjects often used to evaluate the effectiveness and/or harms of interventions in treatment, prevention or diagnosis of disease**
- **Efficiency –**
  - **(in physics) ratio of useful work to the energy supplied to it**
  - **In clinical trials, getting valid and reliable answers to important questions with the least amount of resources**
    - **Does not mean putting patients at risk because of less valid data**
    - **Lower sample size does not mean less work = MORE planning**

32

Lower Sample Size = More Planning

- ***" Clinical trials with small numbers of participants, however, must address broad sets of issues different from those that must be addressed in trials with large numbers of participants. It is in those circumstances of trials with small sample sizes that approaches to optimization of the study design and data interpretation pose greater challenges."***
  - **Small Clinical Trials: Issues and Challenges, Institute of Medicine, 2001 http://www.nap.edu/catalog/10078.html p. ix**

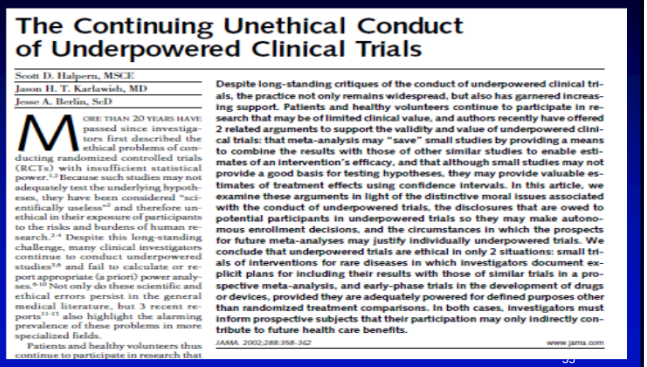
33

Definitions

- What is a “large” or “small” clinical trial?
- IOM defines a “large” trial as one that has adequate sample size to answer the primary research question = “large enough”
- A trial with very few participants may still have adequate statistical power e.g. if effect size is large
- Balance between exposing research subjects to potential harms of experimental interventions with obtaining valid answers

34

Underpowered Studies and Ethics



The Continuing Unethical Conduct of Underpowered Clinical Trials

Scott D. Halpern, MSCE
Jason H. T. Karlawish, MD
Jesse A. Berlin, ScD

MORE THAN 20 YEARS HAVE passed since investigators first described the ethical problems of conducting randomized controlled trials (RCTs) with insufficient statistical power.[1-3] Because such studies may not adequately test the underlying hypotheses, they have been considered "scientifically useless"[2] and therefore unethical in their exposure of participants to the risks and burdens of human research.[2-4] Despite this long-standing challenge, many clinical investigators continue to conduct underpowered studies[5,6] and fail to calculate or report appropriate (a priori) power analyses.[6-10] Not only do these scientific and ethical errors persist in the general medical literature, but 3 recent reports[11-13] also highlight the alarming prevalence of these problems in more specialized fields.

Patients and healthy volunteers thus continue to participate in research that

Despite long-standing critiques of the conduct of underpowered clinical trials, the practice not only remains widespread, but also has garnered increasing support. Patients and healthy volunteers continue to participate in research that may be of limited clinical value, and authors recently have offered 2 related arguments to support the validity and value of underpowered clinical trials: that meta-analysis may "save" small studies by providing a means to combine the results with those of other similar studies to enable estimates of an intervention's efficacy, and that although small studies may not provide a good basis for testing hypotheses, they may provide valuable estimates of treatment effects using confidence intervals. In this article, we examine these arguments in light of the distinctive moral issues associated with the conduct of underpowered trials, the disclosures that are owed to potential participants in underpowered trials so they may make autonomous enrollment decisions, and the circumstances in which the prospects for future meta-analyses may justify individually underpowered trials. We conclude that underpowered trials are ethical in only 2 situations: small trials of interventions for rare diseases in which investigators document explicit plans for including their results with those of similar trials in a prospective meta-analysis, and early-phase trials in the development of drugs or devices, provided they are adequately powered for defined purposes other than randomized treatment comparisons. In both cases, investigators must inform prospective subjects that their participation may only indirectly contribute to future health care benefits.

JAMA. 2002;288:358-362 www.jama.com

35

Underpowered Studies and Ethics

- ***“A proposed study that cannot answer the question being asked because the necessary sample size cannot be attained should not be conducted on ethical grounds. That is, it is unacceptable to expose patients or research participants to harms, even inconveniences, if there is no prospect that useful and potentially generalizable information will result from the study.”***
  - Small Clinical Trials: Issues and Challenges, Institute of Medicine, 2001 http://www.nap.edu/catalog/10078.html p. 14

36

Small Clinical Trials – Last Resort

- *"The importance of conducting small clinical trials only when there are no alternatives cannot be overemphasized. The committee is not encouraging the use of small clinical trials, but, rather provides advice on strategies that should be considered in the design and analysis of small clinical trials when the opportunity to perform a randomized clinical trial with adequate statistical power is not possible. In doing so, it recognizes that small clinical trials frequently need to be viewed as part of a continuing process of data collection. Thus, for some trials it might be impossible to definitively answer a research question with a high degree of confidence. In those cases, perhaps the best that one can do is assess the next set of questions to be asked."*
  - **Small Clinical Trials: Issues and Challenges, Institute of Medicine, 2001 http://www.nap.edu/catalog/10078.html p. 10-11.**

37

Concerns About Small Clinical Trials

- **Small numbers increase variability and leave much to chance**
- **Statistically significant outcomes may not be generalizable (only apply to circumstances in trial)**
- **Too many variables to assess cause and effect**
- **Only able to discern gross effects and limited ability to analyze covariates**
- **Incapable of identifying adverse events**
  - **Small Clinical Trials: Issues and Challenges, Institute of Medicine, 2001 http://www.nap.edu/catalog/10078.html p. 15**

38

Situations Where Smaller Clinical Trials Justifiable

- **Rare diseases**
- **Unique study populations (e.g. astronauts)**
- **Individually tailored therapies**
- **Environments that are isolated**
- **Emergency situations**
- **Public health urgencies**
  - **Small Clinical Trials: Issues and Challenges, Institute of Medicine, 2001 http://www.nap.edu/catalog/10078.html p. 6.**

39

Small vs Efficient

- While small clinical trials are a last resort, efficient clinical trials are always justifiable
- Different methods to improve efficiency are useful (or not) depending on disease under study and research question/setting

40

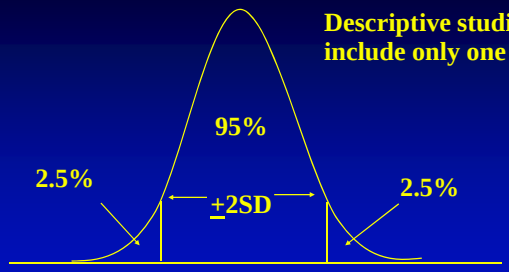
Components of Clinical Studies

- 1. Clear objective of study
- 2. If comparative (rather than descriptive) quantitative comparison with control group
- 3. Select patients for inclusion in study
- 4. If comparative, baseline comparability of groups compared
- 5. Minimizing bias of study
- 6. Well-defined and reliable outcome measures (patient-centered)
- 7. Appropriate statistical analysis

41

Review of Sample Size Considerations

Normal Distribution of a Sample

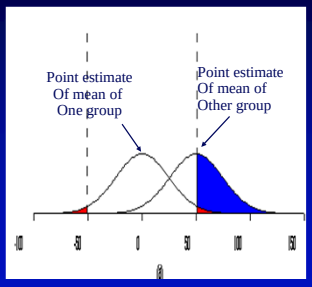


Descriptive studies include only one curve

95% of data will be within 1.96 standard deviations of sample mean for large samples (>30)

43

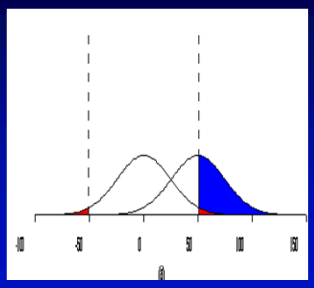
Sample Size – Superiority Trials



- Clinical trials compare average effects in groups of subjects administered intervention to those not administered intervention
- Examine if populations differ by more than chance (for superiority trials)
- Example: Two groups with point estimate for means of 50 and zero, sample size of 12 per group, SD =60

44

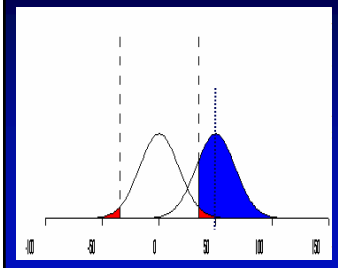
Sample Size – Superiority Trials



- In this example
  - Blue area represents power (in this case 0.497)
  - "critical region" of 0.05 test represented by dashed lines and red area
  - To show difference due to greater than chance, want mean of one curve to be outside of red area

45

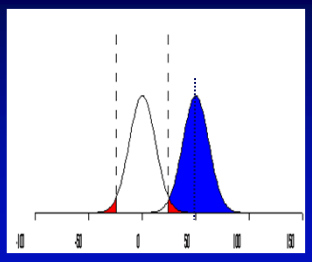
Sample Size



- Sample size in this example is 24 per group
- As sample size increases, overlap between curves decreases (assuming there really is a difference to show)
- Blue area increases = power is 0.80
- Mean of one group now outside of "critical area"
- Notice still a good deal of overlap – only *mean* value is outside critical area

46

Sample Size



- Sample size increased to 48 per group
- Distribution of data narrower and more precise
- Power = 0.98
- Mean well outside of "critical area"

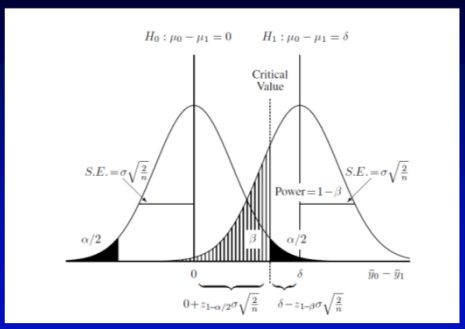
47

Sample Size

- Want to select sample size large enough to show a difference if there is one to detect, but not too large
  - Do not want to expose subjects unnecessarily to harm since this is an experiment evaluating interventions with unknown harm/benefits
  - Use of resources - time, effort and money
- Sample size based on four parameters ("ingredients")
  - Type 1 error- usually specified as 0.05 two sided (0.025 on either side of curve)
  - Type 2 error (1- type 2 error is power) usually specified as 0.10 to 0.20 (power of 80%-90%)
  - Standard deviation of data (variability)
  - Treatment difference – Difference between point estimate of effect for intervention and point estimate for effect with control (delta)

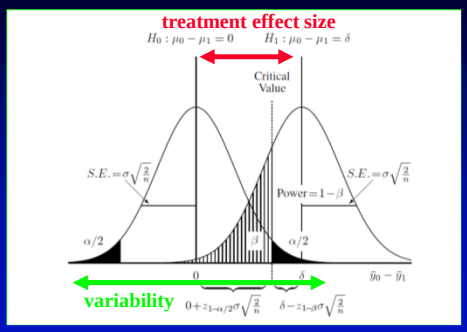
48

Pictorial Representation of Sample Size



49

Pictorial Representation of Sample Size



50

Efficient Clinical Trials

- **Ways to decrease sample size**
  1. **Focused and relevant research question**
  2. **Changing error rates (not suggested)**
  3. **Enhancing effect sizes**
     - **More homogenous populations**
     - **Choosing populations in whom effect size is larger**
     - **Optimizing exposure**
     - **Continuous instead of dichotomous outcomes**
     - **Selection of outcomes**
  4. **Decreasing variability**
     - **More sensitive/specific measures**
     - **Assuring follow-up of enrolled subjects**
     - **Study designs**
       - **Cross-over**
       - **N=1 studies**
       - **Sequential trial designs (e.g. dose-escalation studies)**

51

Efficient Clinical Trials

- Ways to decrease sample size
  1. Focused and relevant research question
  2. Changing error rates (not suggested)
  3. Enhancing effect sizes
     - More homogenous populations
     - Choosing populations in whom effect size is larger
     - Optimizing exposure
     - Continuous instead of dichotomous outcomes
     - Selection of outcomes
  4. Decreasing variability
     - More sensitive/specific measures
     - Assuring follow-up of enrolled subjects
     - Study designs
       - Cross-over
       - N=1 studies
       - Sequential trial designs (e.g. dose-escalation studies)

52

1. Focusing the Research Question

- Research question needs to be one worth answering and of public health importance
- Need to focus question – more questions mean greater sample size or less clear answers = simplify
- "Many trials include measurements to try to figure out why the trial didn't work after it has failed" – the post-mortem on "what the experiment died of"

53

1. Focusing the Research Question

- Sample size is calculated AFTER one decides on a research question
- Starting out with a sample size and working back to "what can I get for this" is not justifiable in terms of choosing unrealistic or clinically meaningless/unachievable effect sizes

54

Efficient Clinical Trials

- Ways to decrease sample size
  1. Focused and relevant research question
  2. Changing error rates (not suggested)
  3. Enhancing effect sizes
     - More homogenous populations
     - Choosing populations in whom effect size is larger
     - Optimizing exposure
     - Continuous instead of dichotomous outcomes
     - Selection of outcomes
  4. Decreasing variability
     - More sensitive/specific measures
     - Assuring follow-up of enrolled subjects
     - Study designs
       - Cross-over
       - N=1 studies
       - Sequential trial designs (e.g. dose-escalation studies)

55

2. Changing Error Rates

- Error rates of type 1 = 0.05 and type 2 = 0.10 or 0.20 are by convention
- But…false positive error rate of 1 in 20 trials is actually a low level of evidence; need justification to deviate from this
- Increasing type 2 error rate increases likelihood of false negative conclusions = spending resources for unclear answers
- Error rates measure random error (by chance) but not bias due to poorly designed study

56

Efficient Clinical Trials

- Ways to decrease sample size
  1. Focused and relevant research question
  2. Changing error rates (not suggested)
  3. Enhancing effect sizes
     - More homogenous populations
     - Choosing populations in whom effect size is larger
     - Optimizing exposure
     - Continuous instead of dichotomous outcomes
     - Selection of outcomes
  4. Decreasing variability
     - More sensitive/specific measures
     - Assuring follow-up of enrolled subjects
     - Study designs
       - Cross-over
       - N=1 studies
       - Sequential trial designs (e.g. dose-escalation studies)

57

Enhancing Effect Sizes

- Trials measure average effects in groups of subjects – "one size fits all" approach may not be correct
- *More homogenous populations* can both decrease variability and increase effect sizes in presence of effect modification
- Effect modification is presence of quantitatively different effect sizes of an intervention based on a baseline variable (e.g. drug is more effective in older people vs younger people)
- Requires knowledge of natural history of disease and evidence from prior trials
- Choosing population in which effect size is larger decreases sample size

58

Examples

- Enrolling subjects in trial in whom effect is expected to be zero
  - Dilutes effect size
  - Ethical issues of exposure to harm for no benefit
- Trastuzumab (Herceptin) in breast cancer
  - Mechanism of action by binding to HER2 proteins in person with specific genetic mutation
  - 20% to 30% of person with breast cancer have this mutation
  - Potentially harmful in those without the mutation

59

Efficient Clinical Trials

- Ways to decrease sample size
  1. Focused and relevant research question
  2. Changing error rates (not suggested)
  3. Enhancing effect sizes
     - More homogenous populations
     - Choosing populations in whom effect size is larger
     - Optimizing exposure
     - Continuous instead of dichotomous outcomes
     - Selection of outcomes
  4. Decreasing variability
     - More sensitive/specific measures
     - Assuring follow-up of enrolled subjects
     - Study designs
       - Cross-over
       - N=1 studies
       - Sequential trial designs (e.g. dose-escalation studies)

60

Optimizing Exposure

- Many trials include only one dose of an intervention with wide inter-individual variabililty in exposure
- Optimize dose based on pre-clinical and early clinical studies – pharmacokinetics and pharmacodynamics
  - Forms a hypothesis to test
  - Not a substitute for clinical trials
- Standardize exposure of interventions - need unblinded third party to do this to maintain blinding in trial

61

Efficient Clinical Trials

- Ways to decrease sample size
  1. Focused and relevant research question
  2. Changing error rates (not suggested)
  3. Enhancing effect sizes
     - More homogenous populations
     - Choosing populations in whom effect size is larger
     - Optimizing exposure
     - Continuous instead of dichotomous outcomes
     - Selection of outcomes
  4. Decreasing variability
     - More sensitive/specific measures
     - Assuring follow-up of enrolled subjects
     - Study designs
       - Cross-over
       - N=1 studies
       - Sequential trial designs (e.g. dose-escalation studies)

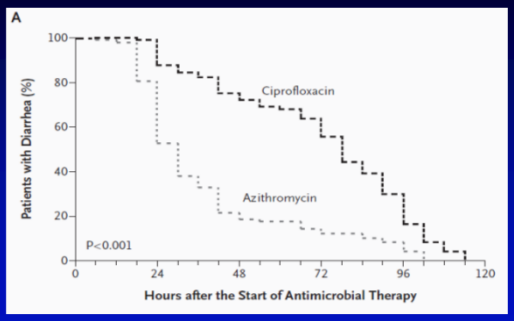
62

Continuous vs Dichotomous Outcomes

- Continuous outcomes have more power to detect differences since uses all the data
- Dichotomous outcomes require
  - 1) categorization – assumes all data in a single category are similarly important which may not be true
  - 2) choosing correct time point to evaluate – if you're wrong, you miss it
- Requires more frequent data capture – patient diaries or phone contact collected in systematic way

63

Time to Event - Cholera



Saha D et al. N Engl J Med 2006;354:2452-62.

64

Efficient Clinical Trials

- Ways to decrease sample size
  1. Focused and relevant research question
  2. Changing error rates (not suggested)
  3. Enhancing effect sizes
     - More homogenous populations
     - Choosing populations in whom effect size is larger
     - Optimizing exposure
     - Continuous instead of dichotomous outcomes
     - Selection of outcomes
  4. Decreasing variability
     - More sensitive/specific measures
     - Assuring follow-up of enrolled subjects
     - Study designs
       - Cross-over
       - N=1 studies
       - Sequential trial designs (e.g. dose-escalation studies)

65

Common or Composite Outcomes

- Choosing more common events or composite outcomes increases number of events and increases power to detect differences
- Some problems with interpretation:
  - Only use when outcomes measured are of similar importance to patients – if driven by less important outcomes may mask inferiority on more important outcomes
  - Does not necessarily imply beneficial effect on all part of a composite

66

Composite Endpoints - Issues

Table I. Structure of data on death and hospitalization (hypothetical data).

	Treatment	
	Active	Control
Died but never hospitalized during follow-up (%)	15	5
Hospitalized and died during follow-up (%)	5	15
Hospitalized, alive at the end of follow-up (%)	20	20
None of the above (%)	60	60

Lubsen J et al. Stat Med 2002;21:2959-70.

67

Efficient Clinical Trials

- **Ways to decrease sample size**
  1. **Focused and relevant research question**
  2. **Changing error rates (not suggested)**
  3. **Enhancing effect sizes**
     - **More homogenous populations**
     - **Choosing populations in whom effect size is larger**
     - **Optimizing exposure**
     - **Continuous instead of dichotomous outcomes**
     - **More common or composite outcomes**
  4. **Decreasing variability**
     - **More sensitive/specific measures**
     - **Assuring follow-up of enrolled subjects**
     - **Study designs**
       - **Cross-over**
       - **N=1 studies**
       - **Sequential trial designs (e.g. dose-escalation studies)**

68

More Sensitive/Specific Measures

- **All measurements composed of the true value plus some associated error**
- **Error can be of two types:**
  - **Random – by chance alone**
  - **Systematic bias – based on individual biases and interpretations**
- **Decreasing error ("noise") in relation to true measure ("signal") allows smaller sample size**

69

Non-Standardized Measures and Error

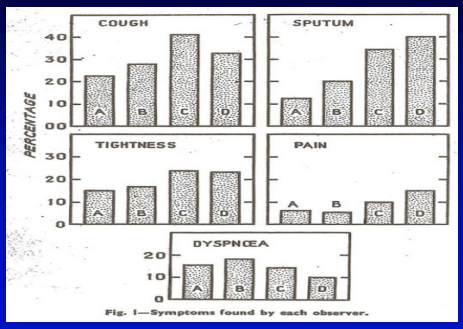


Fig. 1—Symptoms found by each observer.

Cochrane A et al Lancet 1951:1007-9.

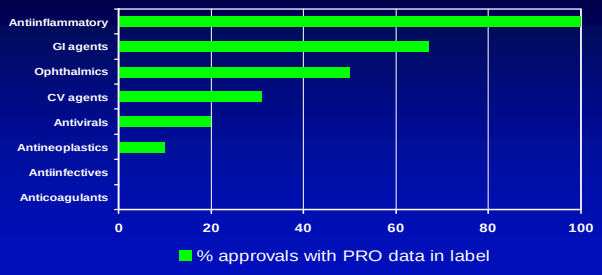
70

Non-Standardized Measures and Error

- Moertel CG and Hanley JA, Effect of measuring error on the results of therapeutic trials in advanced cancer. Cancer 1976;38:388-94.
- 16 oncologists asked to measure 12 simulated tumor masses, two pairs of which were identical in size
- Allowed assessment of 64 measurements by same investigator and 1920 comparisons by different investigators
- 25% "reduction" as "response" for identical size masses = 19% response rate by same investigator, 25% between investigators (measurement error alone)

71

NMEs with PROs in approved labeling, by therapeutic category, '97-'01 *(Willke et al, CCT 2004)*



Note: Only therapeutic classes with at least 9 approvals are included.

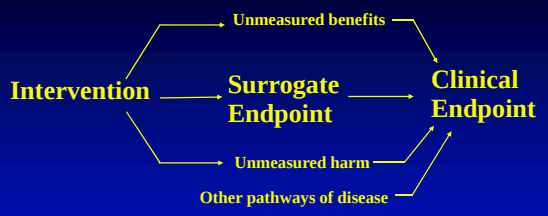
72

Surrogate Endpoints

- Researchers often suggest biomarkers as "surrogate endpoints" in clinical trials – NOT direct measures of patient benefit
- Idea originally was to decrease follow-up time in chronic diseases – keep following subjects to validate biomarker (e.g. viral load in HIV/AIDS, cholesterol in stroke/MI prevention)
- Why use a surrogate in an acute disease when one can measure actual clinical outcomes?
- Surrogate as part of composite outcomes drive the entire outcome since more common

73

Strengths and Limitations



Reasons why surrogate may not accurately predict clinical outcomes:

- unmeasured harms caused by intervention
- unmeasured benefits caused by intervention
- other mechanisms of disease other than those affected by intervention
- issues with measuring surrogate
- issues with measuring clinical outcomes

74

Efficient Clinical Trials

- Ways to decrease sample size
  1. Focused and relevant research question
  2. Changing error rates (not suggested)
  3. Enhancing effect sizes
     - More homogenous populations
     - Choosing populations in whom effect size is larger
     - Optimizing exposure
     - Continuous instead of dichotomous outcomes
     - More common or composite outcomes
  4. Decreasing variability
     - More sensitive/specific measures
     - Assuring follow-up of enrolled subjects
     - Study designs
       - Cross-over
       - N=1 studies
       - Sequential trial designs (e.g. dose-escalation studies)

75

Follow-Up and Missing Data

- Enrolling subjects and then losing them to follow-up or missing data (failure to collect or losing it) results in effort for no gain
- Requires planning on part of researchers and work during the trial to make it as easy as possible for research subjects to return
  - Phone calls and reminders
  - Transportation
  - Home visits
- Subjects who don't follow protocol are not "missing" and should be included in Intent to Treat (ITT) analyses

76

Efficient Clinical Trials

- Ways to decrease sample size
  1. Focused and relevant research question
  2. Changing error rates (not suggested)
  3. Enhancing effect sizes
     - More homogenous populations
     - Choosing populations in whom effect size is larger
     - Optimizing exposure
     - Continuous instead of dichotomous outcomes
     - More common or composite outcomes
  4. Decreasing variability
     - More sensitive/specific measures
     - Assuring follow-up of enrolled subjects
     - Study designs
       - Cross-over
       - N=1 studies
       - Sequential trial designs (e.g. dose-escalation studies)

77

Study Designs

- Certain kinds of study designs can decrease variability and thereby decrease sample size
- Cross-over and n=1 trials both use subjects as their own controls
- Randomize subjects to receive one intervention or the other (sometimes with wash out period in between)

78

Study Designs

- Limitations of cross-over and n=1 trials
  - Most useful in chronic illnessess with stable course of disease
  - If effects carried over from one period treatment to the next then can bias study results
  - "Period effect" for instance in seasonal diseases
  - Most useful in diseases where treatment effect is rapid onset or rapid cessation when intervention stopped

79

Study Designs

- Sequential trial designs often used in Phase 1 dose escalation clinical trials
- Based on pre-defining what level of adverse events or not will allow progression to the next dose
- Makes decision based on data acquired during trial

80

Study Designs

- "Adaptive" designs - word that describes a variety of changes in trial design based on data accumulated during the trial
- Sequential trial dose escalation design is one form but more challenging when modifying other variables like the outcome measure
- Advantage = more subjects assigned to more succesful treatment
- Disadvantage = heterogeneity of subjects based on important risk factors which change as trial progresses introduces bias over time

81

A caveat….

- All the modification we have discussed apply to SUPERIORITY trials
- Non-inferiority is misnomer – does not mean "not inferior" as to show an intervention is "equal" or "not inferior by any amount requires showing superiority
- Biases which trend results toward no difference in a superiority trial (like too small a sample size) result in false positive conclusions in non-inferiority trials
- Non-inferiority does not answer question of added benefit of new interventions; use only in selected situations (e.g improved convenience)

82

Non-inferiority trials are unethical because they disregard patients' interests

Silvio Garattini, Vittorio Bertele'

Equivalence trials[1] have been widely used to assess new drugs, but have recently lost ground to a non-inferiority design. This type of trial is usually accepted by regulatory authorities for approval of new drugs or new indications, although the US Food and Drugs Administration has raised some concerns.[2] In this paper, we argue that the scientific community should ban non-inferiority and equivalence trials because they are unethical, whatever measures are taken to prevent their methodological pitfalls and inappropriate interpretation of results.[3,4] Exceptions might exist, but we could not identify a situation in which patients can justifiably be entered into a trial that will not provide them with any advantage.

but not to the extent that it is recognised as such. For example, if the non-inferiority limit is set at 7·5%, an increase in the incidence of serious events or deaths—say 7% instead of the 5% currently established for the comparator—is not seen as large enough to mark a difference between the new and the control drug. The new drug will therefore be considered non-inferior to the old drug, even if in 1000 patients treated with the former, there could be 20 more deaths than with the latter.

These arguments also apply to equivalence trials, which aim to prove similarity of a new drug to the comparator, since true equivalence is theoretical and is difficult to

Lancet 2007; 370: 1875–77
Published Online
October 23, 2007
DOI:10.1016/S0140-6736(07)61604-3

Mario Negri Institute for Pharmacological Research, Milan, Italy (Prof S Garattini MD, V Bertele' MD)

Correspondence to:
Dr Vittorio Bertele', Mario Negri Institute for Pharmacological Research, Milan, 20156, Italy
bertele@marionegri.it

83

Conclusions

- Developing an efficient trial starts with planning and a good research question
- Question comes first, sample size second
- Various methods to increase effect sizes and decrease variability, when applied in the correct setting, can provide valid and reliable answers to important public health questions
- For some diseases, developing the tools (better outcome measures, better data on natural history) is a good start to better trials

84